Link-Tracing Sampling: Estimating the Size of a Hidden Population in Presence of Heterogeneous Nomination Probabilities

Martín H. Félix-Medina* Pedro E. Monjardin† Aida N. Aceves-Castro ‡

Abstract
Félix-Medina and Thompson (Jour. Off. Stat., 2004) proposed a variant of link-tracing sampling to sample hidden and/or hard-to-detect human populations such as drug users and sex workers. In their variant, an initial sample of venues is selected and the people found in the sampled venues are asked to nominate other members of the population to be included in the sample. Those authors derived maximum likelihood estimators of the population size under the assumption that the probability that a person is nominated (nomination probability) does not depend on the nominee (homogeneity assumption). In this work we extend their research into the case of heterogeneous nomination probabilities and derive both unconditional and conditional maximum likelihood estimators of the population size. In addition, we propose both Wald and profile likelihood confidence intervals for the size of the population. The results of simulations studies carried out by us show that in presence of heterogeneous probabilities the proposed estimators perform reasonably well, whereas the estimators derived under the homogeneity assumption perform badly. The simulation results also indicate that the proposed Wald and profile likelihood confidence intervals have relatively low performance.

Key Words: chain-referral sampling, design-based estimator, maximum likelihood estimator, profile likelihood confidence interval, snowball sampling, Wald confidence interval

1. Introduction

Link-tracing sampling (LTS), also known as snowball sampling or chain referral sampling, has been proposed for sampling hidden or hard-to-detect populations, such as drug users, sex workers, HIV infected people and undocumented workers. In this method an initial sample of members of the target population is selected and the people in the initial sample are asked to nominate other members of the population to be included in the sample. The nominated people that are not in the initial sample might be asked to nominate other persons, and the process might continue in this way until a specified stopping rule is satisfied.

Félix-Medina and Thompson (2004) proposed a variant of LTS in which the initial sample is a simple random sample without replacement (SRSWOR) of sites selected from a sampling frame that could not cover the whole population. The sites are venues where the members of the population might be found with high probabilities, such as public parks, bars and street blocks. The members of the population who belong to a sampled site are identified and they are asked to nominate other members of the population. In order to obtain a maximum likelihood estimator (MLE) of the size of the population, those authors assumed that the probability that a person is nominated by any person in a particular sampled site, which we will

*Escuela de Ciencias Físico-Matemáticas, Universidad Autónoma de Sinaloa, Ciudad Universitaria, Culiacán Sinaloa, México
†Escuela de Ciencias Físico-Matemáticas, Universidad Autónoma de Sinaloa, Ciudad Universitaria, Culiacán Sinaloa, México
‡Escuela de Ciencias Físico-Matemáticas, Universidad Autónoma de Sinaloa, Ciudad Universitaria, Culiacán Sinaloa, México

call nomination probability, depends on the site, but not on the nominated person, that is, they assumed homogeneous nomination probabilities.

In this paper, we extend the work by Félix-Medina and Thompson (2004) to the case in which the nomination probabilities depend on the nominees, that is, we assume heterogeneous nomination probabilities. The structure of this paper is as follows. In Section 2 we introduce the LTS variant proposed by Félix-Medina and Thompson (2004). In Section 3 we derive two types of MLE's of the population size: unconditional and conditional MLE's. In Section 4 we present Wald and profile likelihood confidence intervals for the population size. In Section 5 we present the results of two simulation studies, and finally, in Section 6 we present some conclusions and suggestions for future research.

2. Sampling design

In this work we consider the LTS design proposed by Félix-Medina and Thompson (2004). Thus, let U be a finite population of an unknown number τ of people. We assume that a portion U_1 of U is covered by a sampling frame of N sites $A_1, \ldots, A_N$, where the members of the population can be found with high probability. We suppose that we have a criterion that allows us to assign a person in U_1 to only one site in the frame. Notice that we are not assuming that a person could not be found in different sites, but that we are able to assign him or her to only one site, for instance, the site where he or she spends most of his or her time. Let M_i denote the number of members of the population that belong to the site A_i, $i = 1, \ldots, N$. From the previous assumption it follows that the number of people in U_1 is $\tau_1 = \sum_1^N M_i$ and the number of people in the portion $U_2 = U - U_1$ of U that is not covered by the frame is $\tau_2 = \tau - \tau_1$.

The sampling design is as follows. A SRSWOR S_A of n sites $A_1, \ldots, A_n$ is selected from the frame. The M_i members of the population that belong to the sampled site A_i are identified, $i = 1, \ldots, n$. Let S_0 be the set of people in the initial sample. Notice that the size of S_0 is $M = \sum_1^n M_i$. The people in each sampled site are asked to nominate other members of the population. We will say that a person is nominated by a site if any of the people in that site nominates him or her. We suppose that the nominations made by people in different sites are independent. For each nominated person we record the portion U_1 or U_2 of the population where he or she was located, and the sites that nominated him or her.

3. Maximum likelihood estimators

3.1 Probability models

As in Félix-Medina and Thompson (2004), we will suppose that the numbers $M_1, \ldots,$ M_N of people who belong to the sites $A_1, \ldots, A_N$ are independent Poisson random variables with mean λ_1. Therefore, the joint conditional distribution of the numbers of members $M_1, \ldots, M_n$ in the sampled sites and the number of people $\tau_1 - M$ who are not in the initial sample is multinomial with probability mass function (pmf):

$$f(m_1, \ldots, m_n, \tau_1 - m) = \frac{\tau_1!}{\prod_1^n (\tau_1 - m)!} \left(\frac{1}{N}\right)^m \left(1 - \frac{n}{N}\right)^{\tau_1 - m}. \qquad (1)$$

To model the nomination process we will define random variables $X_{ij}^{(k)}$'s by $X_{ij}^{(k)} =$ 1 if person j in $U_k - A_i$ is nominated by site A_i and $X_{ij}^{(k)} = 0$ otherwise, $i =$

$1, \ldots, n$; $j = 1, \ldots, \tau_k$. We will suppose that the $X_{ij}^{(k)}$'s are independent Bernoulli random variables with means $p_{ij}^{(k)}$'s, where the nomination probability $p_{ij}^{(k)}$ satisfies the following Rash model:

$$p_{ij}^{(k)} = \Pr(X_{ij}^{(k)} = 1) = \frac{\exp(\alpha_i^{(k)} + \beta_j^{(k)})}{1 + \exp(\alpha_i^{(k)} + \beta_j^{(k)})}, \quad j \in U_k - A_i; \quad i = 1, \ldots, n. \qquad (2)$$

It is worth noting that this model was considered by Coull and Agresti (1999) in the context of multiple capture-recapture sampling. In this model $\alpha_i^{(k)}$ is a fixed (not random) effect that represents the potential of the site A_i to nominate a person in $U_k - A_i$, and $\beta_j^{(k)}$ is a random effect that represents the propensity of the person $j \in U_k$ to be nominated. We will suppose that $\beta_j^{(k)}$ is normally distributed with mean 0 and unknown variance σ_k^2 and that these variables are independent. The parameter σ_k^2 determines the degree of heterogeneity of the $p_{ij}^{(k)}$'s, that is, great values of σ_k^2 imply high degrees of heterogeneity.

3.2 Likelihood function

To construct the likelihood function we will first define the concept of "pattern of nomination". Thus, let $\Omega = \{1, \ldots, n\}$ and let ω be a subset of Ω. We will say that a person has pattern of nomination ω if that person is nominated only by each one of the sites A_i with $i \in \omega$. For instance, if $\omega = \{1, 3, 4\}$, a person has pattern of nomination ω if he or she is nominated only by the sites A_1, A_3 and A_4. For convenience we will say that a person that is not nominated by any of the sites has pattern of nomination $\omega = \emptyset$ (the empty set). Notice that the set of people in $U_k - S_0$ can be partitioned into the set of all the subsets ω of Ω. From the assumed Rash model for the $p_{ij}^{(k)}$'s we have that the probability that a randomly selected person in $U_k - S_0$ has pattern of nomination ω is

$$\pi_\omega^{(k)}(\sigma_k, \alpha^{(k)}) = \int \prod_{i=1}^{n} \frac{e^{x_{\omega i}(\alpha_i^{(k)} + \sigma_k z)}}{1 + e^{(\alpha_i^{(k)} + \sigma_k z)}} \phi(z) dz,$$

where $x_{\omega i} = 1$ if $i \in \omega$ and $x_{\omega i} = 0$ otherwise, $\alpha^{(k)} = (\alpha_1^{(k)}, \ldots, \alpha_1^{(k)})$ and $\phi(\cdot)$ is the probability density function (pdf) of the standard normal distribution [N(0,1)].

As in Coull and Agresti (1999), instead of using $\pi_\omega^{(k)}(\sigma_k, \alpha^{(k)})$ in the likelihood function we will use its Gaussian quadrature approximation $\tilde{\pi}_\omega^{(k)}(\sigma_k, \alpha^{(k)})$ given by

$$\tilde{\pi}_\omega^{(k)}(\sigma_k, \alpha^{(k)}) = \sum_{t=1}^{q} \prod_{i=1}^{n} \frac{e^{x_{\omega i}(\alpha_i^{(k)} + \sigma_k z_t)}}{1 + e^{(\alpha_i^{(k)} + \sigma_k z_t)}} \nu_t,$$

where q is a fixed constant and $\{z_t\}$ and $\{\nu_t\}$ are obtained from tables.

The easiest way of constructing the likelihood function is to factorize it into different factors. One factor is the one associated with the selection procedure of the initial sample S_0, which is given by the multinomial distribution (1), that is

$$L_{MULT}(\tau_1) \propto \frac{\tau_1!}{(\tau_1 - m)!}(1 - n/N)^{\tau_1 - m}.$$

Two other factors are the ones associated with the nomination processes of the people in $U_1 - S_0$ and in U_2. These factors are obtained by using the fact that the set of people in each of these regions can be partitioned into the set of all different patterns of nomination. Therefore, these factors are given by the multinomial distributions of the respective frequencies (numbers of people) of the different patterns of nomination, that is,

$$L_1(\tau_1,\sigma_1,\alpha^{(1)}) \propto \frac{(\tau_1-m)!}{(\tau_1-m-r_1)!}\prod_{\omega\subset\Omega-\emptyset}[\tilde{\pi}^{(1)}_{\omega}(\sigma_1,\alpha^{(1)})]^{r^{(1)}_{\omega}}\,[\tilde{\pi}^{(1)}_{\emptyset}(\sigma_1,\alpha^{(1)})]^{\tau_1-m-r_1}$$

and

$$L_2(\tau_2,\sigma_2,\alpha^{(2)}) \propto \frac{\tau_2!}{(\tau_2-r_2)!}\prod_{\omega\subset\Omega-\emptyset}[\tilde{\pi}^{(2)}_{\omega}(\sigma_2,\alpha^{(2)})]^{r^{(2)}_{\omega}}\,[\tilde{\pi}^{(2)}_{\emptyset}(\sigma_2,\alpha^{(2)})]^{\tau_2-r_2},$$

where r_k is the number of distinct nominated persons in $U_k - S_0$ and $r^{(k)}_{\omega}$ is the number of persons in $U_k - S_0$ that have the pattern of nomination ω, $k = 1, 2$.

The last factor $L_0(\sigma_1, \alpha^{(1)})$ is the one associated with the nomination process of the people in S_0 and it is given by

$$L_0(\sigma_1,\alpha^{(1)}) \propto \prod_{i=1}^{n} L_{A_i}(\sigma_1,\alpha^{(1)}_{-i}),$$

where $\alpha^{(1)}_{-i} = (\alpha^{(1)}_1,\ldots,\alpha^{(1)}_{i-1},\alpha^{(1)}_{i+1},\ldots,\alpha^{(1)}_n)$ and $L_{A_i}(\sigma_1,\alpha^{(1)}_{-i})$ is the factor associated with the nomination process of the people in $A_i \in S_A$, given by the multinomial distribution of the frequencies (numbers of people) of the different patterns of nomination in which the people in A_i can be partitioned. Thus

$$L_{A_i}(\sigma_1,\alpha^{(1)}_{-i}) \propto \prod_{\omega\subset\Omega_i-\emptyset}[\tilde{\pi}^{(A_i)}_{\omega}(\sigma_1,\alpha^{(1)})]^{r^{(A_i)}_{\omega}}\,[\tilde{\pi}^{(A_i)}_{\emptyset}(\sigma_1,\alpha^{(1)})]^{m_i-r^{(A_i)}_{\omega}},$$

where $\Omega_i = \Omega - \{i\}$, $r^{(A_i)}_{\omega}$ is the number of people in A_i that has the pattern of nomination $\omega \subset \Omega_i$ and

$$\tilde{\pi}^{(A_i)}_{\omega}(\sigma_1,\alpha^{(1)}_{-i}) = \sum_{t=1}^{q}\prod_{i'\neq i}^{n}\frac{e^{x_{\omega i'}(\alpha^{(1)}_{i'}+\sigma_1 z_t)}}{1+e^{(\alpha^{(1)}_{i'}+\sigma_1 z_t)}}\nu_t,$$

where $x_{\omega i'} = 1$ if $i' \in \omega$ and $x_{\omega i'} = 0$ otherwise, is the Gaussian quadrature approximation to the probability that a person in A_i has pattern of nomination $\omega \subset \Omega_i$, $i = 1, \ldots, n$.

From the previous results we have that the maximum likelihood function is given by

$$L(\tau_1,\tau_2,\sigma_1,\sigma_2,\alpha^{(1)},\alpha^{(2)}) = L_{(1)}(\tau_1,\sigma_1,\alpha^{(1)})L_{(2)}(\tau_2,\sigma_2,\alpha^{(2)}),$$

where

$$L_{(1)}(\tau_1,\sigma_1,\alpha^{(1)}) = L_{MULT}(\tau_1)L_1(\tau_1,\sigma_1,\alpha^{(1)})L_0(\sigma_1,\alpha^{(1)}) \quad \text{and}$$

$$L_{(2)}(\tau_2,\sigma_2,\alpha^{(2)}) = L_2(\tau_2,\sigma_2,\alpha^{(2)}).$$

3.3 Unconditional maximum likelihood estimators

Numerical maximization of L with respect to the parameters yields the unconditional MLEs $\hat{\tau}_k$, $\hat{\sigma}_k$ and $\hat{\alpha}^{(k)}$ of τ_k, σ_k and $\alpha^{(k)}$, $k = 1, 2$. The MLE of $\tau = \tau_1 + \tau_2$ is $\hat{\tau} = \hat{\tau}_1 + \hat{\tau}_2$.

3.4 Conditional maximum likelihood estimators

From a numerical point of view a simpler approach to estimate the τ's is the one proposed by Sanathan (1972) based on conditional maximum likelihood estimation. The idea is to factorize the multinomial distributions of the frequencies $r_\omega^{(k)}$ of the different patterns of nomination as follows:

$$\begin{aligned}
L_1(\tau_1,\sigma_1,\alpha^{(1)}) &\propto f(\{r_\omega^{(1)}\}|m,\tau_1,\sigma_1,\alpha^{(1)}) = f(\{r_\omega^{(1)}\}|r_1,\sigma_1,\alpha^{(1)})f(r_1|m,\tau_1,\sigma_1,\alpha^{(1)})\\
&\propto \prod_{\omega\subset\Omega-\emptyset}\left[\frac{\tilde{\pi}_\omega^{(1)}(\sigma_1,\alpha^{(1)})}{1-\tilde{\pi}_\emptyset^{(1)}(\sigma_1,\alpha^{(1)})}\right]^{r_\omega^{(1)}}\\
&\quad\times\frac{(\tau_1-m)!}{(\tau_1-m-r_1)!}[1-\tilde{\pi}_\emptyset^{(1)}(\sigma_1,\alpha^{(1)})]^{r_1}[\tilde{\pi}_\emptyset^{(1)}(\sigma_1,\alpha^{(1)})]^{\tau_1-m-r_1}\\
&= L_{11}(\sigma_1,\alpha^{(1)})L_{12}(\tau_1,\sigma_1,\alpha^{(1)})
\end{aligned}$$

and

$$\begin{aligned}
L_2(\tau_2,\sigma_2,\alpha^{(2)}) &\propto f(\{r_\omega^{(2)}\}|\tau_2,\sigma_2,\alpha^{(2)}) = f(\{r_\omega^{(2)}\}|r_2,\sigma_2,\alpha^{(2)})f(r_2|\tau_2,\sigma_2,\alpha^{(2)})\\
&\propto \prod_{\omega\subset\Omega-\emptyset}\left[\frac{\tilde{\pi}_\omega^{(2)}(\sigma_2,\alpha^{(2)})}{1-\tilde{\pi}_\emptyset^{(2)}(\sigma_2,\alpha^{(2)})}\right]^{r_\omega^{(2)}}\\
&\quad\frac{\tau_2!}{(\tau_2-r_2)!}[1-\tilde{\pi}_\emptyset^{(2)}(\sigma_2,\alpha^{(2)})]^{r_2}[\tilde{\pi}_\emptyset^{(2)}(\sigma_2,\alpha^{(2)})]^{\tau_2-r_2}\\
&= L_{21}(\sigma_2,\alpha^{(2)})L_{22}(\tau_2,\sigma_2,\alpha^{(2)}).
\end{aligned}$$

Notice that in each case the first factor $L_{k1}(\sigma_k,\alpha^{(k)})$ is proportional to the pmf of the multinomial distribution with parameter of size R_k and vector of probabilities $(\{\tilde{\pi}_\omega^{(k)}/[1-\tilde{\pi}_\emptyset^{(k)}]\}_{\omega\in\Omega-\emptyset})$, which does not depend on τ_k. Notice also that the second factors $L_{12}(\tau_1,\sigma_1,\alpha^{(1)})$ and $L_{22}(\tau_2,\sigma_2,\alpha^{(2)})$ are proportional to the pmf's of the $\mathrm{Bin}(\tau_1-m,1-\tilde{\pi}_\emptyset^{(1)})$ and $\mathrm{Bin}(\tau_2,1-\tilde{\pi}_\emptyset^{(2)})$, respectively, where $\mathrm{Bin}(\tau,\pi)$ denotes the Binomial distribution with parameter of size τ and probability π.

Numerical maximizations of

$$L_{11}(\sigma_1,\alpha^{(1)})L_0(\sigma_1,\alpha^{(1)}) \quad\text{and}\quad L_{21}(\sigma_2,\alpha^{(2)}) \tag{3}$$

with respect to $(\sigma_1,\alpha^{(1)})$ and $(\sigma_2,\alpha^{(2)})$, respectively, give the conditional MLE's $\breve{\sigma}_k$ and $\breve{\alpha}^{(k)}$ of σ_k and $\alpha^{(k)}$, $k=1,2$. Notice that the factors in (3) do not depend on τ_k, $k=1,2$. Finally, plugging the estimates $\breve{\sigma}_k$ and $\breve{\alpha}^{(k)}$ into the factors of the likelihood function that depend on τ_k, and maximizing these factors, that is, maximizing

$$L_{12}(\tau_1,\breve{\sigma}_1,\breve{\alpha}^{(1)})L_{MULT}(\tau_1) \quad\text{and}\quad L_{22}(\tau_2,\breve{\sigma}_2,\breve{\alpha}^{(2)}),$$

with respect to τ_1 and τ_2, respectively, yield the following conditional MLE's of these parameters:

$$\begin{aligned}
\breve{\tau}_1 &= \frac{M+R_1}{1-(1-n/N)\breve{\tilde{\pi}}_\emptyset^{(1)}(\breve{\sigma}_1,\breve{\alpha}^{(1)})} \quad\text{and}\\
\breve{\tau}_2 &= \frac{R_2}{1-\breve{\tilde{\pi}}_\emptyset^{(2)}(\breve{\sigma}_2,\breve{\alpha}^{(2)})}.
\end{aligned}$$

A conditional MLE of τ is $\breve{\tau}=\breve{\tau}_1+\breve{\tau}_2$.

4. Confidence intervals for the population sizes

4.1 Wald confidence intervals

Here we will suppose that if the numbers of nominees R_k, $k = 1, 2$, are large, then the standard normal distribution is a reasonable approximation to the distribution of $(\hat{\tau}_k - \tau_k)/\sqrt{V(\hat{\tau}_k)}$, where $\hat{\tau}_k$ denotes either the unconditional or the conditional MLE of τ_k and $V(\hat{\tau}_k)$ denotes its variance. Under this assumption, an approximate $100(1-\alpha)\%$ Wald confidence interval for τ_k is

$$\hat{\tau}_k \pm z_{1-\alpha/2}\sqrt{\hat{V}(\hat{\tau}_k)}, \quad k = 1, 2,$$

where $z_{1-\alpha/2}$ denotes the $(1-\alpha/2)$th quantile point of the $N(0,1)$ and $\hat{V}(\hat{\tau}_k)$ denotes an estimator of the variance of $\hat{\tau}_k$.

By using Sanathan's (1972) approach to derive an asymptotic approximation to the variance of the unconditional or conditional MLE of the parameter of size of a multinomial distribution, we have that if R_k, $k = 1, 2$, and M_i, $i = 1, \dots, N$, are large, then an estimator of the variance of τ_k is

$$\hat{V}(\hat{\tau}_k) = \hat{\tau}_k[\hat{D}_k - \hat{B}_k'\hat{A}_k^{-1}\hat{B}_k]^{-1},$$

where $\hat{A}_k = [\hat{a}_{ij}^{(k)}]_{n+1,n+1}$,

$$\hat{a}_{ij}^{(1)} = \left(1 - \frac{n}{N}\right)\sum_{\omega\subseteq\Omega}\frac{1}{\hat{\tilde{\pi}}_\omega^{(1)}}\left(\frac{\partial\hat{\tilde{\pi}}_\omega^{(1)}}{\partial\hat{\theta}_i^{(1)}}\right)\left(\frac{\partial\hat{\tilde{\pi}}_w^{(1)}}{\partial\hat{\theta}_j^{(1)}}\right) + \frac{1}{N}\sum_{\omega\subseteq\Omega_i}\frac{1}{\hat{\tilde{\pi}}_\omega^{(A_i)}}\left(\frac{\partial\hat{\tilde{\pi}}_\omega^{(A_i)}}{\partial\hat{\theta}_i^{(1)}}\right)\left(\frac{\partial\hat{\tilde{\pi}}_\omega^{(A_i)}}{\partial\hat{\theta}_j^{(1)}}\right),$$

$$\hat{a}_{ij}^{(2)} = \sum_{\omega\subseteq\Omega}\frac{1}{\hat{\tilde{\pi}}_\omega^{(2)}}\left(\frac{\partial\hat{\tilde{\pi}}_\omega^{(2)}}{\partial\hat{\theta}_i^{(2)}}\right)\left(\frac{\partial\hat{\tilde{\pi}}_w^{(2)}}{\partial\hat{\theta}_j^{(2)}}\right),$$

$$\hat{B}_k = (\hat{b}_1^{(k)}, \dots, \hat{b}_{n+1}^{(k)})', \quad \hat{b}_j^{(k)} = -(\partial\hat{\tilde{\pi}}_\emptyset^{(k)}/\partial\hat{\theta}_j^{(k)})/\hat{\tilde{\pi}}_\emptyset^{(k)}, \quad j = 1, \dots, n+1, \quad k = 1, 2,$$

$$D_1 = [1 - (1 - n/N)\hat{\tilde{\pi}}_\emptyset^{(1)}]/[(1 - n/N)\hat{\tilde{\pi}}_\emptyset^{(1)}], \quad D_2 = [1 - \hat{\tilde{\pi}}_\emptyset^{(2)}]/\hat{\tilde{\pi}}_\emptyset^{(2)} \quad \text{and}$$

$$\hat{\theta}^{(k)} = (\hat{\theta}_1^{(k)}, \dots, \hat{\theta}_n^{(k)}, \hat{\theta}_{n+1}^{(k)}) = (\hat{\alpha}_1^{(k)}, \dots, \hat{\alpha}_n^{(k)}, \hat{\sigma}_k), \quad k = 1, 2.$$

A variance estimator of $\hat{\tau}$ is $\hat{V}(\hat{\tau}) = \hat{V}(\hat{\tau}_1) + \hat{V}(\hat{\tau}_2)$. It is worth noting that these expressions for the variance estimators are valid either if $\hat{\tau}_k$, $\hat{\sigma}_k$ and $\hat{\alpha}^{(k)}$ represent the unconditional or the conditional MLE's.

Félix-Medina and Thompson (2004) showed that their MLE's of the population sizes derived under the homogeneity assumption are robust to deviations from the assumed Poisson distribution of the M_i's, but that the "model-based" estimators of the variances are not; therefore, we expect that our proposed estimators of the population sizes and the "previous model-based" estimators of their variances behave as the ones proposed by those authors. Thus, we have also derived "partly design-based" estimators of the variances of the $\hat{\tau}$'s, which are obtained by replacing the multinomial distribution of the M_i's by the distribution that was used to select the initial sample of sites; however, because these estimators still use the assumptions about the model that describe the nomination process, they are not completely "design-based", but only "partly design-based". The purpose of deriving these estimators is that they should be robust to deviations from the assumed Poisson distribution of the M_i's. The "partly design-based" estimator $\hat{V}_D(\hat{\tau}_1)$ of the variance of $\hat{\tau}_1$ is given by the expression for $\hat{V}(\hat{\tau}_1)$, but computing $\hat{D}_1$ as follows

$$D_1 = \frac{n}{\hat{\tau}_1(1 - n/N)}S_M^2 + \frac{1 - \hat{\tilde{\pi}}_\emptyset^{(1)}}{(1 - n/N)\hat{\tilde{\pi}}_\emptyset^{(1)}},$$

where

$$S_M^2 = \sum_1^n (M_i - \bar{M})^2/(n-1) \quad \text{and} \quad \bar{M} = \sum_1^n M_i/n,$$

are the sample variance and sample mean of the M_i's. The partly design-based estimator $\hat{V}_D(\hat{\tau}_2)$ of the variance of $\hat{\tau}_2$ is the same as the model-based estimator of its variance, and the design-based estimator of the variance of $\hat{\tau}$ is $\hat{V}_D(\hat{\tau}) = \hat{V}_D(\hat{\tau}_1) + \hat{V}_D(\hat{\tau}_2)$.

4.2 Profile likelihood intervals

In the context of multiple capture-recapture sampling, several authors such as Evans et al. (1996) and Gimenez et al. (2005), have indicated that Wald confidence intervals perform poorly when the sample sizes are not large. The reasons of their poor performance are (i) biases in the estimates of the population size; (ii) biases in the estimates of the variances and (iii) asymmetries in the distributions of the estimators of the population sizes. They have also pointed out that a more robust method for constructing intervals than Wald method is the one based on the asymptotic chi-square (χ^2) distribution of the generalized likelihood ratio test. The confidence intervals obtained by this method are called profile likelihood confidence intervals (PLCI). In our case, an approximate $100(1-\alpha)\%$ PLCI for τ_k is given by

$$\{\tau_k : -2\ln[\Lambda(\tau_k)] \leq \chi^2_{1,1-\alpha}\},$$

where

$$\Lambda(\tau_k) = \max_{\sigma_k,\, \alpha^{(k)}} L_{(k)}(\tau_k, \sigma_k, \alpha^{(k)})/L_{(k)}(\hat{\tau}_k, \hat{\sigma}_k, \hat{\alpha}^{(k)}), \tag{4}$$

$\hat{\tau}_k$, $\hat{\sigma}_k$ and $\hat{\alpha}^{(k)}$ are the unconditional MLE's of τ_k, σ_k and $\alpha^{(k)}$, $k = 1, 2$, and $\chi^2_{1,1-\alpha}$ is the $(1-\alpha)$th quantile point of the χ^2 distribution with 1 degree of freedom.

To obtain a PLCI for $\tau = \tau_1 + \tau_2$, we use the equality $\tau_1 = \tau - \tau_2$ and consider the likelihood function of τ, τ_2, σ_1, σ_2, $\alpha^{(1)}$ and $\alpha^{(2)}$, which is given by

$$L(\tau, \tau_2, \sigma_1, \sigma_2, \alpha^{(1)}, \alpha^{(2)}) = L_{(1)}(\tau - \tau_2, \sigma_1, \alpha^{(1)})L_{(2)}(\tau_2, \sigma_2, \alpha^{(2)}).$$

Then, an approximate $100(1-\alpha)\%$ PLCI for τ is given by (4), but using $L(\tau, \tau_2, \sigma_1, \sigma_2, \alpha^{(1)}, \alpha^{(2)})$ instead of $L_{(k)}(\tau_k, \sigma_k, \alpha^{(k)})$ and computing the numerator of $\Lambda(\tau)$ by maximizing $L(\tau, \tau_2, \sigma_1, \sigma_2, \alpha^{(1)}, \alpha^{(2)})$ with respect to τ_2, σ_1, σ_2, $\alpha^{(1)}$ and $\alpha^{(2)}$.

As in the case of the model-based Wald confidence intervals, we do not expect that these model-based PLCI's for τ_1 and τ be robust to deviations from the assumed Poisson distribution of the M_i's. Therefore, we have considered adjusted PLCI's that take into account extra Poisson variation of the M_i's. These intervals are constructed as the model-based ones, but replacing the $(1-\alpha)$th quantile point $\chi^2_{1,1-\alpha}$ of the χ^2 distribution with one degree of freedom by $(S_M^2/\bar{M})\chi^2_{1,1-\alpha}$.

5. Monte Carlo studies

We carried out two simulation studies to explore the performance of the proposed estimators of the population sizes, estimators of their variances and confidence intervals, as well as to compare their performance with that of the corresponding ones derived under the homogeneity assumption and which were proposed by Félix-Medina and Thompson (2004). Because the computation of both the proposed Wald confidence intervals and the proposed PLCI's require a lot of computing time, we

Table 1: Characteristics of the artificial populations generated for the first study.

Population I	Population II
$N = 250$	$N = 250$
$M_i \sim$ Poisson	$M_i \sim$ Neg. binomial
$E(M_i) = 7.2$	$E(M_i) = 7.2$
$V(M_i) = 7.2$	$V(M_i) = 24.5$
$\tau_1 = 1926$	$\tau_1 = 1774$
$\tau_2 = 750$	$\tau_2 = 750$
$\tau = 2676$	$\tau = 2524$
$\tau_1/\tau = 0.72$	$\tau_1/\tau = 0.70$
$\alpha_i^{(k)} = \frac{-6.5}{M_i^{1/4}+0.001}$	$\alpha_i^{(k)} = \frac{-6.5}{M_i^{1/4}+0.001}$
$\beta_j^{(k)} \sim N(0, 0.75)$	$\beta_j^{(k)} \sim N(0, 0.75)$

needed to carry out two simulation studies. In the first one we analyzed the performance of the estimators of the population sizes by using populations and initial samples of relatively large sizes. In the second study we observed the performance of the estimators of the variances and confidence intervals by using populations and initial samples of small size.

5.1 First simulation study

We considered two finite populations of $N = 250$ m_i-values. In Population I the values were generated from a Poisson distribution with mean 7.2, whereas in Population II from a Negative binomial distribution with mean 7.2 and variance 24.5 In Table 1 are displayed the characteristics of each population. Notice that in both populations the value of τ_2 was set to 750. The values of the nomination probabilities $p_{ij}^{(k)}$'s were obtained by means of Rash model (2), where the values of the $\alpha_i^{(k)}$'s and $\beta_j^{(k)}$'s were obtained as is indicated on Table 1. The values of these parameters were set so that the average value of the $p_{ij}^{(k)}$'s was about 0.025, and since we used an initial sample of size $n = 25$, the probability of nominating a person by any of the sampled sites was about 0.46 in Population I and 0.43 in Population II.

The simulation study was carried out by replicating $r = 500$ times the following procedure. From each population of $N = 250$ values of M_i's a SRSWOR of $n = 25$ values was selected. From the i-th selected value, $i =, 1, \ldots, n$, the values of $X_{ij}^{(1)}$, $j = 1, \ldots, \tau_1 - M_i$, and $X_{ij'}^{(2)}$, $j' = 1, \ldots, \tau_2$, were obtained from Bernoulli distributions with means $p_{ij}^{(1)}$, $j = 1, \ldots, \tau_1 - M_i$, and $p_{ij'}^{(2)}$, $j' = 1, \ldots, \tau_2$, respectively. These data on the M_i's and the $X_{ij}^{(k)}$'s were used to compute the estimates of the population sizes. In this study we considered the estimators $\tilde{\tau}_1$, $\tilde{\tau}_2$ and $\tilde{\tau}$ proposed by Félix-Medina and Thompson (2004) and derived under the homogeneity assumption; the unconditional $\hat{\tau}_1$, $\hat{\tau}_2$ and $\hat{\tau}$ and the conditional $\check{\tau}_1$, $\check{\tau}_2$ and $\check{\tau}$ MLE's proposed in this work. The performance of an estimator $\hat{\tau}$, say, of τ was evaluated by means of its relative bias (r-bias) and the square root of its relative mean square error (r-mse) defined by r-bias $= \sum_1^r(\hat{\tau}_i - \tau)/(r\tau)$ and $\sqrt{\text{r-mse}} = \sqrt{\sum_1^r(\hat{\tau}_i - \tau)^2/(r\tau^2)}$, where $\hat{\tau}_i$ was the value of $\hat{\tau}$ obtained in the i-th trial.

Table 2: Means, relative biases and square roots of relative mean square errors of estimators of τ_1, τ_2 and τ.

	Population I			Population II		
	$n = 25$ $\bar{M} = 192.5$ $\bar{R}^{(1)} = 725.0$ $\bar{R}^{(2)} = 306.1$			$n = 25$ $\bar{M} = 177.8$ $\bar{R}^{(1)} = 612.1$ $\bar{R}^{(2)} = 296.7$		
Estimator	Mean	r-bias	$\sqrt{\text{r-mse}}$	Mean	r-bias	$\sqrt{\text{r-mse}}$
$\tilde{\tau}_1$	1530.9	-0.21	0.21	1424.5	-0.20	0.20
$\tilde{\tau}_2$	518.4	-0.31	0.31	510.1	-0.32	0.32
$\tilde{\tau}$	2049.3	-0.23	0.24	1934.6	-0.23	0.24
$\hat{\tau}_1$	1938.4	0.01	0.05	1788.5	0.01	0.07
$\hat{\tau}_2$	760.0	0.01	0.19	822.7	0.10	0.33
$\hat{\tau}$	2698.4	0.01	0.06	2611.2	0.04	0.11
$\check{\tau}_1$	1955.8	0.02	0.06	1770.1	-0.00	0.07
$\check{\tau}_2$	755.3	0.01	0.17	731.7	-0.02	0.16
$\check{\tau}$	2711.2	0.01	0.07	2501.8	-0.01	0.07

Notes: $\tilde{\tau}_1$, $\tilde{\tau}_2$ and $\tilde{\tau}$ MLEs proposed by Félix- Medina and Thompson (2004) derived under a homogeneous nomination probabilities model; $\hat{\tau}_1$, $\hat{\tau}_2$ and $\hat{\tau}$ and $\check{\tau}_1$, $\check{\tau}_2$ and $\check{\tau}$ unconditional and conditional MLEs. Results based on 500 samples.

The results of the study are shown on Table 2. We can see that the estimators $\tilde{\tau}_1$, $\tilde{\tau}_2$ and $\tilde{\tau}$ that do not take into account the heterogeneity of the $p_{ij}^{(k)}$'s had problems of biases and these increased the r-mse's of the estimators significantly. On the other hand, our proposed unconditional and conditional MLE's performed satisfactorily, although the conditional estimators performed better than the unconditional estimators. Notice that the performance of the estimators was not affected by the distribution of the M_i's. This means that they were robust to deviations from the Poisson distribution. Notice also that both the unconditional and the conditional MLE's of τ_2 showed some problems of instability (the values of $\sqrt{\text{r-mse}}$ are between 0.16 and 0.33).

5.2 Second simulation study

This study was carried out using the same procedure as that used in the first one. Thus, we generated two artificial populations whose characteristics are shown on Table 3. It is worth noting that the values of the $\alpha_i^{(k)}$'s and $\beta_j^{(k)}$'s were set so that the average value of the $p_{ij}^{(k)}$'s was about 0.09. In this study, the size of the initial sample was set to $n = 10$; therefore, the probability of nominating a person by any of the sampled sites was about 0.62.

The 95% confidence intervals that were considered in this study were the partly design-based Wald intervals based on the estimators $\tilde{\tau}_1$, $\tilde{\tau}_2$ and $\tilde{\tau}$ proposed by Félix-Medina and Thompson (2004) who derived them under the homogeneity assumption; both the model-based and the partly design-based Wald confidence intervals based on the unconditional $\hat{\tau}_1$, $\hat{\tau}_2$ and $\hat{\tau}$ and on the conditional $\check{\tau}_1$, $\check{\tau}_2$ and $\check{\tau}$ MLE's proposed in this work, and both the PLCI's and the adjusted for extra Poisson variation PLCI's also proposed in this work. It is worth noting that the variance

Table 3: Characteristics of the artificial populations generated for the second study.

Population I	Population II
$N = 100$	$N = 100$
$M_i \sim$ Poisson	$M_i \sim$ Neg. binomial
$E(M_i) = 7.2$	$E(M_i) = 7.0$
$V(M_i) = 7.2$	$V(M_i) = 23.3$
$\tau_1 = 725$	$\tau_1 = 716$
$\tau_2 = 500$	$\tau_2 = 500$
$\tau = 1225$	$\tau = 1216$
$\tau_1/\tau = 0.59$	$\tau_1/\tau = 0.59$
$\alpha_i^{(k)} = \frac{-4.0}{M_i^{1/4}+0.001}$	$\alpha_i^{(k)} = \frac{-4.0}{M_i^{1/4}+0.001}$
$\beta_j^{(k)} \sim N(0, 0.75)$	$\beta_j^{(k)} \sim N(0, 0.75)$

estimators that were used to compute Wald intervals were also evaluated. The performance of a variance estimator was evaluated by using the same criteria as those used for an estimator of the population size. The performance of a confidence interval for τ, say, was evaluated by its coverage probability (CP) and its relative length (r-length) defined as the proportion of trials in which τ is inside the interval and the average length of the interval divided by τ, respectively. The evaluation of the performance of a Wald interval was based on 500 samples, whereas that of a PLCI was based on 100 samples.

The results of the study are shown on Tables 4 and 5. From the results we can see that the estimators of the variances of the estimators of the population sizes obtained under the homogeneity assumption had large negative biases. These serious subestimations of the variances along with the biases of the estimators of the population sizes substantially deteriorated the coverage probabilities of the corresponding Wald intervals.

With respect to the proposed estimators of the variances of the unconditional and conditional MLE's of the population sizes, both the proposed model-based and the proposed partly design-based estimators of the variances had tolerable performance in Population I. The estimators of the variances of $\hat{\tau}_2$ and $\check{\tau}_2$ had relatively large variability which increased the variability of the estimators of the variances of $\hat{\tau}$ and $\check{\tau}$. However, the corresponding Wald confidence intervals performed acceptably well, with the exception of the intervals for τ_2 which were relatively long. In the case of Population II, the estimators of the variances of the unconditional and conditional MLE's of the τ's did not perform well. This means that they, including the partly design-based, were not robust to deviations from the assumed Poisson distribution of the M_i's. The estimators of the variances of $\hat{\tau}_1$ and $\check{\tau}_1$ had large biases. The estimators of the variances of $\hat{\tau}_2$ and $\check{\tau}_2$ were highly unstable and this instability affected the stability of the estimators of the variances of $\hat{\tau}$ and $\check{\tau}$. Wald confidence intervals for $\hat{\tau}_1$ and $\check{\tau}_1$ had relatively low coverage probabilities. Wald intervals for τ_2 were long. However, Wald confidence intervals for τ, the most important parameter, performed well.

With respect to the profile likelihood intervals, in Population I, they had relatively low coverage probabilities. In Population II, the coverage probability of the PLCI for τ_1 was very low because this interval was sensitive to the deviation from

Table 4: Population I. Relative biases and square roots of relative mean square errors of variance estimators and coverage probabilities and relative lengths of 95% confidence intervals for τ_1, τ_2 and τ.

Variance estimator	r-bias	$\sqrt{\text{r-mse}}$	Confidence interval	CP	r-length
$\tilde{V}_D(\tilde{\tau}_1)$	-.29	.31	$\tilde{\tau}_1 \pm 1.96\sqrt{\tilde{V}_D(\tilde{\tau}_1)}$	0.00	.11
$\tilde{V}_D(\tilde{\tau}_2)$	-.27	.33	$\tilde{\tau}_2 \pm 1.96\sqrt{\tilde{V}_D(\tilde{\tau}_2)}$	0.00	.13
$\tilde{V}_D(\tilde{\tau})$	-.36	.38	$\tilde{\tau}_{\pm}1.96\sqrt{\tilde{V}_D(\tilde{\tau})}$	0.00	.08
$\hat{V}(\hat{\tau}_1)$	-.03	.17	$\hat{\tau}_1 \pm 1.96\sqrt{\hat{V}(\hat{\tau}_1)}$	0.92	.24
$\tilde{V}(\hat{\tau}_2)$	.08	.94	$\hat{\tau}_2 \pm 1.96\sqrt{\hat{V}(\hat{\tau}_2)}$	0.94	.52
$\hat{V}(\hat{\tau})$	.13	.71	$\hat{\tau} \pm 1.96\sqrt{\hat{V}(\hat{\tau})}$	0.93	.26
$\hat{V}_D(\hat{\tau}_1)$	.02	.24	$\hat{\tau}_1 \pm 1.96\sqrt{\hat{V}_D(\hat{\tau}_1)}$	0.93	.24
$\tilde{V}_D(\hat{\tau}_2)$	.08	.94	$\hat{\tau}_2 \pm 1.96\sqrt{\hat{V}_D(\hat{\tau}_2)}$	0.94	.52
$\hat{V}_D(\hat{\tau})$	.15	.72	$\hat{\tau} \pm 1.96\sqrt{\hat{V}_D(\hat{\tau})}$	0.93	.26
$\check{V}(\check{\tau}_1)$	-.25	.31	$\check{\tau}_1 \pm 1.96\sqrt{\check{V}(\check{\tau}_1)}$	0.91	.24
$\tilde{V}(\check{\tau}_2)$	.09	.96	$\check{\tau}_2 \pm 1.96\sqrt{\check{V}(\check{\tau}_2)}$	0.94	.59
$\check{V}(\check{\tau})$	.05	.67	$\check{\tau} \pm 1.96\sqrt{\check{V}(\check{\tau})}$	0.94	.26
$\check{V}_D(\check{\tau}_1)$	-.20	.33	$\check{\tau}_1 \pm 1.96\sqrt{\check{V}_D(\check{\tau}_1)}$	0.91	.25
$\tilde{V}_D(\check{\tau}_2)$	.09	.96	$\check{\tau}_2 \pm 1.96\sqrt{\tilde{V}_D(\check{\tau}_2)}$	0.94	.59
$\check{V}_D(\check{\tau})$	.07	.68	$\check{\tau} \pm 1.96\sqrt{\check{V}_D(\check{\tau})}$	0.94	.27
			PLCI based on $\hat{\tau}_1$	0.90	.21
			PLCI based on $\hat{\tau}_2$	0.89	.54
			PLCI based on $\hat{\tau}$	0.84*	.26*

Notes: $\tilde{V}_D(\tilde{\tau}_1)$, $\tilde{V}_D(\tilde{\tau}_2)$ and $\tilde{V}_D(\tilde{\tau})$ partly design-based estimators of the variances of the MLE's $\tilde{\tau}_1$, $\tilde{\tau}_2$ and $\tilde{\tau}$ proposed by Félix-Medina and Thompson (2004) derived under a homogeneous nomination probabilities model; $\hat{V}(\hat{\tau}_1)$, $\hat{V}(\hat{\tau}_2)$ and $\hat{V}(\hat{\tau})$, and $\hat{V}_D(\hat{\tau}_1)$, $\hat{V}_D(\hat{\tau}_2)$ and $\hat{V}_D(\hat{\tau})$ model-based and partly design-based estimators of the variances of the unconditional MLE's $\hat{\tau}_1$, $\hat{\tau}_2$ and $\hat{\tau}$; $\check{V}(\check{\tau}_1)$, $\check{V}(\check{\tau}_2)$ and $\check{V}(\check{\tau})$, and $\check{V}_D(\check{\tau}_1)$, $\check{V}_D(\check{\tau}_2)$ and $\check{V}_D(\check{\tau})$ model-based and partly design-based estimators of the variances of the conditional MLE's $\check{\tau}_1$, $\check{\tau}_2$ and $\check{\tau}$; PLCI, profile likelihood confidence interval; CP, coverage probability; $n = 10$; $\bar{M} = 72.4$, $\bar{R}^{(1)} = 365.9$ and $\bar{R}^{(2)} = 280.3$. Results for Wald intervals were based on 500 samples, whereas results for PLCI 's were based on 100 samples. Results marked with an * were obtained discarding 5 samples.

Table 5: Population II. Relative biases and square roots of relative mean square errors of variance estimators and coverage probabilities and relative lengths of 95% confidence intervals for τ_1, τ_2 and τ.

Variance estimator	r-bias	$\sqrt{\text{r-mse}}$	Confidence interval	CP	r-length
$\tilde{V}_D(\tilde{\tau}_1)$	-.50	.51	$\tilde{\tau}_1 \pm 1.96\sqrt{\tilde{V}_D(\tilde{\tau}_1)}$	0.00	.11
$\tilde{V}_D(\tilde{\tau}_2)$	-.32	.41	$\tilde{\tau}_2 \pm 1.96\sqrt{\tilde{V}_D(\tilde{\tau}_2)}$	0.00	.14
$\tilde{V}_D(\tilde{\tau})$	-.50	.51	$\tilde{\tau}\pm 1.96\sqrt{\tilde{V}_D(\tilde{\tau})}$	0.00	.09
$\hat{V}(\hat{\tau}_1)$	-.35	.37	$\hat{\tau}_1 \pm 1.96\sqrt{\hat{V}(\hat{\tau}_1)}$	0.89	.24
$\hat{V}(\hat{\tau}_2)$	$.05^4$	1.3^4	$\hat{\tau}_2 \pm 1.96\sqrt{\hat{V}(\hat{\tau}_2)}$	0.92^4	$.63^4$
$\hat{V}(\hat{\tau})$	$.02^4$	$.82^4$	$\hat{\tau} \pm 1.96\sqrt{\hat{V}(\hat{\tau})}$	0.95^4	$.30^4$
$\hat{V}_D(\hat{\tau}_1)$	-.55	.56	$\hat{\tau}_1 \pm 1.96\sqrt{\hat{V}_D(\hat{\tau}_1)}$	0.79	.20
$\hat{V}_D(\hat{\tau}_2)$	$.05^4$	1.3^4	$\hat{\tau}_2 \pm 1.96\sqrt{\hat{V}_D(\hat{\tau}_2)}$	0.92^4	$.63^4$
$\hat{V}_D(\hat{\tau})$	$-.01^4$	$.82^4$	$\hat{\tau} \pm 1.96\sqrt{\hat{V}_D(\hat{\tau})}$	0.92^4	$.29^4$
$\check{V}(\check{\tau}_1)$	-.26	.34	$\check{\tau}_1 \pm 1.96\sqrt{\check{V}(\check{\tau}_1)}$	0.92	.25
$\check{V}(\check{\tau}_2)$	$.24^4$	1.3^4	$\check{\tau}_2 \pm 1.96\sqrt{\check{V}(\check{\tau}_2)}$	0.94^4	$.65^4$
$\check{V}(\check{\tau})$	$.05^4$	$.89^4$	$\check{\tau} \pm 1.96\sqrt{\check{V}(\check{\tau})}$	0.97^4	$.31^4$
$\check{V}_D(\check{\tau}_1)$	-.47	.52	$\check{\tau}_1 \pm 1.96\sqrt{\check{V}_D(\check{\tau}_1)}$	0.84	.21
$\check{V}_D(\check{\tau}_2)$	$.24^4$	1.3^4	$\check{\tau}_2 \pm 1.96\sqrt{\check{V}_D(\check{\tau}_2)}$	0.94^4	$.65^4$
$\check{V}_D(\check{\tau})$	$-.50^4$	1.0^4	$\check{\tau} \pm 1.96\sqrt{\check{V}_D(\check{\tau})}$	0.96^4	$.30^4$
			PLCI based on $\hat{\tau}_1$	0.76^2	$.21^2$
			PLCI based on $\hat{\tau}_2$	0.95^4	$.58^4$
			PLCI based on $\hat{\tau}$	0.92^{12}	$.33^{12}$
			Adjusted PLCI based on $\hat{\tau}_1$	0.93	.39
			Adjusted PLCI based on $\hat{\tau}_2$	0.95^4	$.58^4$
			Adjusted PLCI based on $\hat{\tau}$	0.98	.49

Notes: $\tilde{V}_D(\tilde{\tau}_1)$, $\tilde{V}_D(\tilde{\tau}_2)$ and $\tilde{V}_D(\tilde{\tau})$ partly design-based estimators of the variances of the MLE's $\tilde{\tau}_1$, $\tilde{\tau}_2$ and $\tilde{\tau}$ proposed by Félix-Medina and Thompson (2004) derived under a homogeneous nomination probabilities model; $\hat{V}(\hat{\tau}_1)$, $\hat{V}(\hat{\tau}_2)$ and $\hat{V}(\hat{\tau})$, and $\hat{V}_D(\hat{\tau}_1)$, $\hat{V}_D(\hat{\tau}_2)$ and $\hat{V}_D(\hat{\tau})$ model-based and partly design-based estimators of the variances of the unconditional MLE's $\hat{\tau}_1$, $\hat{\tau}_2$ and $\hat{\tau}$; $\check{V}(\check{\tau}_1)$, $\check{V}(\check{\tau}_2)$ and $\check{V}(\check{\tau})$, and $\check{V}_D(\check{\tau}_1)$, $\check{V}_D(\check{\tau}_2)$ and $\check{V}_D(\check{\tau})$ model-based and partly design-based estimators of the variances of the conditional MLE's $\check{\tau}_1$, $\check{\tau}_2$ and $\check{\tau}$; PLCI, profile likelihood confidence interval; CP, coverage probability; $n = 10$; $\bar{M} = 71.2$, $\bar{R}^{(1)} = 358.9$ and $\bar{R}^{(2)} = 260.0$. Results for Wald intervals were based on 500 samples, whereas results for PLCI 's were based on 100 samples. Results marked with a superscript were obtained discarding the number of samples indicated by the superscript.

the Poisson distribution of the M_i's. However, the adjusted PLCI's, that take into account the extra Poisson variation of the M_i's, showed good coverage probabilities.

It is worth noting that for some samples the values of the estimators of the variances of $\hat{\tau}_2$ and $\check{\tau}_2$ were huge. These huge values caused that the values of the estimators of the variances of $\hat{\tau}$ and $\check{\tau}$ were also huge. The problem was that the computation of each estimator of the variance of $\hat{\tau}_2$ or $\check{\tau}_2$ requires to solve a system of linear equations and for some samples the matrices of coefficients were ill-conditioned. As is indicated in Table 5, we discarded these cases. In addition, PLCI's could not be computed for some samples. The problem was that the upper limit of the interval is obtained as the largest root of the equation $-2\ln[\Lambda(\tau_k)] = \chi^2_{1,1-\alpha}$, and for some samples that root was huge. These cases were also discarded.

6. Conclusions and suggestions for further research

Even though the simulation studies carried out in this research were not extensive, according to the results we can say that in presence of heterogeneous nomination probabilities the two types of proposed MLE's of the population sizes perform acceptably regardless of the distribution of the M_i's. On the other hand, the MLE's derived under the homogeneity assumption perform badly: they have large biases that increase their mean square errors substantially. The proposed variance estimators perform tolerably well when the Poisson distribution of the M_i's is satisfied, but they do not perform well when that assumption is not satisfied. This implies that they are not robust to deviations from the Poisson distribution. Similarly, Wald confidence intervals perform acceptably well when the assumption of Poisson distribution is satisfied, but they do not perform well when that assumption is not satisfied, with the exception of the intervals for τ which perform acceptably well. The results about the performance of the profile likelihood intervals do not allow us to reach any conclusion. For instance, low coverage probabilities were obtained when the assumption of Poisson distribution of the M_i's is satisfied, whereas coverage probabilities close to 0.95 were obtained (with the exception of the coverage probability of the interval for τ_1) when that assumption was not satisfied. Therefore, further research is required to reach conclusions about the performance of this type of interval.

It is worth noting that not only the performance of the PLCI's needs further research, but the performance of the proposed estimators of the population sizes, estimators of their variances and Wald intervals also requires further research since the Monte Carlo studies carried out in this research were very limited. Another problem that needs further research is the implementation or development of efficient numerical methods to compute the proposed point and interval estimators. Their computation requires a lot of calculations and consequently is very computing time consuming. A possibility could be to implement the method to compute PLCI's proposed by Venzon and Moolgavkar (1988).

REFERENCES

Coull, B. A., and Agresti, A. (1999), "The use of mixed logit models to reflect heterogeneity in capture-recapture studies," *Biometrics*, 55, 294-301.

Evans, M. A., Kim, H.-M., and O'Brien, T. E. (1996), "An application of profile-likelihood based confidence interval to capture-recapture estimators," *Journal of Agricultural, Biological and Environmental Statistics*, 1, 131-140.

Félix-Medina, M. H., and Thompson, S. K. (2004), "Combining cluster sampling and link-tracing sampling to estimate the size of hidden populations," *Journal of Official Statistics*, 20, 19-38.

Gimenes, O., Choquet, R., Lamor, R., Scofield, P., Fletcher, D., Lebreton, J.-D., and Pradel, R. (2005), Efficient profile-likelihood confidence intervals for capture-recapture models," *Journal of Agricultural, Biological and Environmental Statistics*, 10, 1-13.

Sanathan, L. (1972), "Estimating the size of a multinomial population," *Annals of Mathematical Statistics*, 43, 142-152.

Venzon, D. J., and Moolgavkar, S. H. (1988), "A method for computing profile-likelihood-based confidence intervals," *Applied Statistics*, 37, 87-94.